Joint Scheduling and Resource Allocation in OFDMA Downlink Systems via ACK/NAK Feedback

Rohit Aggarwal, C. Emre Koksal, and Philip Schniter

Dept. of ECE, The Ohio State University, Columbus, OH 43210.

Email: {aggarwar,koksal,schniter}@ece.osu.edu

Abstract

In this paper, we consider the problem of joint scheduling and resource allocation in the OFDMA downlink, with the goal of maximizing an expected long-term goodput-based utility subject to an instantaneous sum-power constraint, and where the feedback to the base station consists only of ACK/NAKs from recently scheduled users. We first establish that the optimal solution is a partially observable Markov decision process (POMDP), which is impractical to implement. In response, we propose a greedy approach to joint scheduling and resource allocation that maintains a posterior channel distribution for every user, and has only polynomial complexity. For frequency-selective channels with Markov time-variation, we then outline a recursive method to update the channel posteriors, based on the ACK/NAK feedback, that is made computationally efficient through the use of particle filtering. To gauge the performance of our greedy approach relative to that of the optimal POMDP, we derive a POMDP performance upper-bound. Numerical experiments show that, for slowly fading channels, the performance of our greedy scheme is relatively close to the upper bound, and much better than fixed-power random user scheduling (FP-RUS), despite its relatively low complexity.

Keywords: OFDMA downlink, scheduling and resource allocation, ACK/NAK feedback, particle filters.

I. INTRODUCTION

In the downlink of a wireless orthogonal frequency division multiple access (OFDMA) system, the base station (BS) must deliver data to a set of users whose channels may vary in both time and frequency. Since bandwidth and power resources are limited, data delivery must be carried out efficiently, e.g., by pairing users with strong subchannels and by distributing power across users in the most effective manner. Often, the BS must also adhere to per-user quality-of-service (QoS) constraints. Overall, the BS faces the challenging problem of jointly scheduling users across subchannels, optimizing their modulation-and-coding schemes, and allocating a limited power resource to maximize some function of per-user throughputs.

The OFDMA scheduling-and-resource-allocation problem has been addressed in a number of studies that assume the availability of perfect channel state information (CSI) at the BS (e.g., [1]–[7]). In practice, however, it is difficult for the BS to maintain perfect CSI (for all users and all subchannels), since CSI is most easily obtained at the user terminals, and the bandwidth available for feedback of CSI to the BS is scarce. Hence, practical resource allocation schemes use some form of limited feedback [8], such as quantized channel gains.

In this work, we consider the exclusive use of ACK/NAK feedback, as provided by the automatic repeat request (ARQ) [9] mechanism present in most wireless downlinks. We assume standard ARQ,[1] where every scheduled user provides the BS with either an acknowledgment (ACK), if the most recent data packet has been correctly decoded, or a negative acknowledgment (NAK), if not. Although ACK/NAKs do not provide direct information about the state of the channel, they do provide *relative* information about channel quality that can be used for the purpose of transmitter adaptation (e.g., [10], [11]). For example, if an NAK was received for a particular packet, then it is likely that the subchannel's signal-to-noise ratio (SNR) was below that required to support the transmission rate used for that packet. We consider the *exclusive* use ACK/NAK feedback provided by the link layer, because this allows us to completely avoid *any additional* feedback, such as feedback about quantized channel gains.

There are interesting implications to the use of (quantized) *error-rate* feedback (like ACK/NAK) for transmitter adaptation, as opposed to quantized channel-state feedback. With error-rate feed-

[1] The approach we develop in this paper could be easily extended to other forms of link-layer feedback, e.g., Type-I and Type-II Hybrid ARQ. For simplicity and ease of exposition, however, we consider only standard ARQ.

back, the transmission parameters applied at a given time-slot affect not only the throughput for that slot, but also the corresponding feedback, which will impact the quality of future transmitter-CSI, and thus future throughput. For example, if the transmission parameters are chosen to maximize only the instantaneous throughput, e.g., by scheduling those users that the BS believes are currently best, then little will be learned about the changing states of other user channels, implying that future scheduling decisions will be compromised. On the other hand, if the BS schedules not-recently-scheduled users solely for the purpose of probing their channels, then instantaneous throughput will be compromised. Thus, when using error-rate feedback, the BS must navigate the classic tradeoff between exploitation and exploration [12].

In this work, we propose a scheme whereby the BS uses ACK/NAK feedback to maintain a posterior channel distribution for every user and, from these distributions, performs simultaneous user subchannel-scheduling, power-allocation, and rate-selection. In doing so, the BS aims to maximize an expected, long-term, generic *utility* criterion that is a function of the per-user/channel/rate goodputs. Our use of a generic utility-based criterion allows us to handle, e.g., sum-capacity maximization, throughput maximization under practical modulation-and-coding schemes, and throughput-based pricing (e.g., [13]–[15]), as discussed in the sequel. To this end, we exploit our recent work [16], which offers an efficient near-optimal scheme for utility-based OFDMA resource allocation under distributional CSI. Our use of ACK/NAK-feedback, however, makes our problem considerably more complicated than the one considered in [16]. For example, as we show in the sequel, the optimal solution to our expected long-term utility-maximization problem is a *partially observable Markov decision process* (POMDP) that would involve the solution of many mixed-integer optimization problems during each time-slot. Due to the impracticality of the POMDP solution, we instead consider (suboptimal) *greedy* utility-maximization schemes. As justification for this approach, we first establish that the optimal utility maximization strategy would itself be greedy if the BS had perfect CSI for all user-subchannel combinations. Moreover, we establish that the performance of this perfect-CSI (greedy) scheme upper-bounds the optimal ACK/NAK-feedback-based (POMDP) scheme. We then propose a novel, greedy utility-maximization scheme whose performance is shown (via the upper bound) to be close to optimal. Finally, due to the computational demands of tracking the posterior channel distribution for every user, we propose a low-complexity implementation based on particle filtering.

We now describe the relation of our work to the existing literature [17]–[19]. In [17], a *learning-automata*-based user/rate scheduling algorithm was proposed to maximize system throughput based on ACK/NAK feedback while satisfying per-user throughput constraints. While [17] considered a single channel, we consider joint user/rate scheduling and power allocation in a multi-channel OFDMA setting. In [18], a state-space-based approach was taken to jointly schedule users/rates and allocate powers in downlink OFDMA systems under slow-fading channels in the presence of ACK/NAK feedback and imperfect subchannel-gain estimates at the BS. In particular, assuming a discrete channel model, goodput maximization was considered under a target maximum packet-error probability constraint and a sum-power constraint across all time-slots. Its solution led to a POMDP which was solved using a dynamic-program. While the approach in [18] is applicable to only goodput maximization under discrete-state channels, ours is applicable to generic utility maximization problems under continuous-state channels. Furthermore, our approach is based on particle filtering and lends itself to practical implementation. In [19], the user/rate scheduling and power allocation problem in OFDMA systems with quasi-static channels and ACK/NAK feedback was formulated as a Markov Decision Process and an efficient algorithm was proposed to maximize achievable sum-rate while maintaining a target packet-error-rate and a sum-power constraint over a finite time-horizon. Apart from assuming a discrete-state quasi-static channel model, the scope of this work was limited by two other assumptions: *i)* in each time-slot, the BS scheduled only one user across all subchannels for data transmission, and *ii)* all users decoded the broadcasted data-packet and sent ACK/NAK feedback to the BS. In contrast, we consider the scenario where multi-user diversity is efficiently exploited by scheduling different users across different subchannels, and only the scheduled users report ACK/NAK feedback. Furthermore, we consider general utility maximization under continuous-state time-varying channels, and propose a polynomial-complexity joint scheduling and resource allocation scheme with provable performance guarantees.

The rest of the paper is organized as follows. In Section II, we outline the system model and, in Section III, we investigate the optimal scheduling and resource allocation scheme. Due to the implementation complexity of the optimal scheme, we propose a suboptimal greedy scheme in Section IV that maintains posterior channel distributions inferred from the received ACK/NAK feedback. In Section V, we show how these posteriors can be recursively updated via particle filtering. Numerical results are presented in Section VI, and conclusions are stated in Section VII.

II. System Model

We consider a packetized downlink OFDMA system with a pool of K users. During each time slot, the BS (i.e., "controller") transmits packets of data, composed of codewords from a generic signaling scheme, through N OFDMA subchannels (with $N \leqslant K$). Each packet propagates through a fading channel on the way to its intended mobile user, where the fading channel is assumed to be time-invariant over the packet duration, but is allowed to vary across packets in a Markovian manner. Henceforth, we will use "time" when referring to the packet index. At each time-instant, the BS must decide—for each subchannel—which user to schedule, which modulation-and-coding scheme (MCS) to use, and how much power to allocate.

We assume M choices of MCS, where the MCS index $m \in \{1, \dots, M\}$ corresponds to a transmission rate of r_m bits per packet and a packet error rate of the form $\epsilon = a_m e^{-b_m P \gamma}$ under transmit power P and squared subchannel gain (SSG) γ, where a_m and b_m are constants [20]. Let (n, k, m) represent the combination of user k and MCS m over subchannel n. In the sequel, we use $P^t_{n,k,m}$, $\gamma^t_{n,k}$, and $\epsilon^t_{n,k,m}$ to denote—respectively—the power allocated to, the SSG experienced by, and the error rate of the combination (n, k, m) at time t. Additionally, we denote the scheduling decision by $I^t_{n,k,m} \in \{0, 1\}$, where $I^t_{n,k,m} = 1$ indicates that user/rate (k, m) was scheduled on subchannel n at time t, whereas $I^t_{n,k,m} = 0$ indicates otherwise. Since we assume that only one user/rate (k, m) can be scheduled on a given subchannel n at a given time t, we have the "subchannel resource" constraint $\sum_{k,m} I^t_{n,k,m} \leqslant 1$ for all n, t. We also assume a "sum-power constraint" of the form $\sum_{n,k,m} I^t_{n,k,m} P^t_{n,k,m} \leqslant X_{\mathsf{con}}$ for all t.

Our goal in scheduling and resource allocation is to maximize an expected long-term utility criterion that is a function of the per-user/rate/subchannel goodputs, i.e., $\mathrm{E}\left\{\sum_{n,k,m,t} U_{n,k,m}(g^t_{n,k,m})\right\}$. Here, $g^t_{n,k,m}$ denotes the goodput contributed by user k with MCS m on subchannel n at time t, which can be expanded as $g^t_{n,k,m} = I^t_{n,k,m}(1 - \epsilon^t_{n,k,m}) r_m$. Meanwhile, $U_{n,k,m}(\cdot)$ is a generic utility function that we assume (for technical reasons) is twice differentiable, strictly-increasing, and concave, with $U_{n,k,m}(0) < \infty$. We use $U_{n,k,m}(\cdot)$ to transform goodput into other metrics that are more meaningful from the perspective of quality-of-service (QoS), fairness [21], or pricing (e.g., [13]–[15]). For example, to maximize sum-goodput, one would simply use $U_{n,k,m}(x) = x$. To enforce fairness across users, one could instead maximize weighted sum-goodput via $U_{n,k,m}(x) = w_k x$, where $\{w_k\}$ are appropriately chosen user-dependent weights. To maximize

sum capacity, i.e., $\sum_{n,k} I^t_{n,k,1} \log(1 + P^t_{n,k,1}\gamma^t_{n,k})$, one would choose $M = a_1 = b_1 = r_1 = 1$ and $U_{n,k,1}(x) = \log(1 - \log(1 - x))$ for $x \in [0, 1)$. To incorporate user-fairness into capacity maximization, one could instead choose $U_{n,k,1}(\cdot) = w_k \log(1 - \log(1 - x))$, where again $\{w_k\}$ are appropriately chosen user-dependent weights [20].

For each time t, the BS performs scheduling and resource allocation based on posterior distributions on the SSGs $\{\gamma^t_{n,k}\}$ inferred from previously received ACK/NAK feedback. In the sequel, we write the ACK/NAK feedback about the packet transmitted to user k across subchannel n at time t by $f^t_{n,k} \in \{1, 0, \emptyset\}$, where 1 indicates an ACK, 0 indicates a NAK, and $\emptyset$ covers the case that user k was not scheduled on subchannel n at time t. Thus, in the case of an infinite past horizon and a feedback delay of $d \geqslant 1$ packets, the BS would have access to the feedbacks $\{f^\tau_{n,k}\ \forall n, k\}^{t-d}_{\tau=-\infty}$ for time-t scheduling.

III. Optimal Scheduling and Resource Allocation

In this section, we describe the optimal solution to the problem of scheduling and resource allocation over the finite time-horizon $t \in \{1, \ldots, T\}$. For this purpose, some additional notation will be useful. To denote the collection of all time-t scheduling variables $\{I^t_{n,k,m}\}$, we use $\boldsymbol{I}^t \in \{0, 1\}^{NKM}$. To denote the collection of all time-t powers $\{P^t_{n,k,m}\}$, we use $\boldsymbol{P}^t \in [0, \infty)^{NKM}$. To denote the collection of all time-t ACK/NAK feedbacks $\{f^t_{n,k}\}$ we use $\boldsymbol{F}^t \in \{1, 0, \emptyset\}^{NK}$, and to denote the collection of all time-t user-k feedbacks we use $\boldsymbol{f}^t_k \in \{1, 0, \emptyset\}^N$.

For time-t scheduling and resource allocation, the controller has access to the previous feedback $\boldsymbol{F}^{t-d}_{-\infty} \triangleq \{\boldsymbol{F}^{-\infty}, \ldots, \boldsymbol{F}^{t-d}\}$, scheduling decisions $\boldsymbol{I}^{t-d}_{-\infty} \triangleq \{\boldsymbol{I}^{-\infty}, \ldots, \boldsymbol{I}^{t-d}\}$, and power allocations $\boldsymbol{P}^{t-d}_{-\infty} \triangleq \{\boldsymbol{P}^{-\infty}, \ldots, \boldsymbol{P}^{t-d}\}$. It then uses this knowledge to determine the schedule $\boldsymbol{I}^t$ and power allocation $\boldsymbol{p}^t$ maximizing the expected utility of the current and remaining packets:

$$(\boldsymbol{I}^{t,\mathsf{opt}}, \boldsymbol{P}^{t,\mathsf{opt}}) = \operatorname*{arg\,max}_{(\boldsymbol{I}^t,\boldsymbol{P}^t)\in\mathcal{X}} \mathrm{E}\Bigg\{ \sum_{n,k,m} I^t_{n,k,m} U_{n,k,m}\big((1 - a_m e^{-b_m P^t_{n,k,m}\gamma^t_{n,k}}) r_m\big) + \sum_{\tau=t+1}^{T} I^{\tau,\mathsf{opt}}_{n,k,m} U_{n,k,m}\big((1 - a_m e^{-b_m P^{\tau,\mathsf{opt}}_{n,k,m}\gamma^\tau_{n,k}}) r_m\big) \Bigg| \boldsymbol{F}^{t-d}_{-\infty}, \boldsymbol{I}^{t-d}_{-\infty}, \boldsymbol{P}^{t-d}_{-\infty} \Bigg\}, \quad (1)$$

where the domain of $\boldsymbol{I}^t$ is $\mathcal{I} \triangleq \{\boldsymbol{I} \in \{0, 1\}^{NKM} : \sum_{k,m} I_{n,k,m} \leqslant 1\ \forall n\}$, the domain of $\boldsymbol{P}^t$ is $\mathcal{P} \triangleq [0, \infty)^{NKM}$, and $\mathcal{X} \triangleq \{(\boldsymbol{I}, \boldsymbol{P}) \in \mathcal{I} \times \mathcal{P} : \sum_{n,k,m} I_{n,k,m} P_{n,k,m} \leqslant X_{\mathsf{con}}\}$. The expectation in (1) is jointly over the squared subchannel gains (SSGs) $\{\gamma^\tau_{n,k} : \tau = t, \ldots, T, \forall n, \forall k\}$. Using the abbreviations $\tilde{U}^t_{n,k,m}(I_{n,k,m}, P_{n,k,m}) \triangleq I_{n,k,m} U_{n,k,m}\big((1 - a_m e^{-b_m P_{n,k,m}\gamma^t_{n,k}}) r_m\big)$ and $\mathbb{F}^{t-d}_{-\infty} \triangleq$

$\{\boldsymbol{F}_{-\infty}^{t-d}, \boldsymbol{I}_{-\infty}^{t-d}, \boldsymbol{P}_{-\infty}^{t-d}\}$, the optimal expected utility over the remaining packets $\{t, \ldots, T\}$ can be written (for $t \geqslant 0$) as

$$U_{\mathsf{tot}}^{t,\mathsf{opt}}(\mathbb{F}_{-\infty}^{t-d}) \triangleq \mathrm{E}\left\{ \sum_{\tau=t}^{T} \sum_{n,k,m} \tilde{U}_{n,k,m}^{\tau}\big(I_{n,k,m}^{\tau,\mathsf{opt}}, P_{n,k,m}^{\tau,\mathsf{opt}}\big) \,\Big|\, \mathbb{F}_{-\infty}^{t-d} \right\}. \tag{2}$$

For a unit-delay[2] system (i.e. $d = 1$), the following Bellman equation [22] specifies the corresponding finite-horizon dynamic program:

$$\begin{aligned} U_{\mathsf{tot}}^{t,\mathsf{opt}}(\mathbb{F}_{-\infty}^{t-1}) \;=\; \max_{(\boldsymbol{I}^t, \boldsymbol{P}^t) \in \mathcal{X}} \Bigg[&\mathrm{E}\left\{ \sum_{n,k,m} \tilde{U}_{n,k,m}^{t}\big(I_{n,k,m}^{t}, P_{n,k,m}^{t}\big) \,\Big|\, \mathbb{F}_{-\infty}^{t-1} \right\} \\ &+ \mathrm{E}\left\{ U_{\mathsf{tot}}^{t+1,\mathsf{opt}}\big(\mathbb{F}_{-\infty}^{t-1} \cup \{\boldsymbol{F}^t, \boldsymbol{I}^t, \boldsymbol{P}^t\}\big) \,\Big|\, \mathbb{F}_{-\infty}^{t-1} \right\} \Bigg], \end{aligned} \tag{3}$$

where the second expectation is over the feedbacks $\boldsymbol{F}^t$. The solution obtained by solving (1) is typically referred to as a partially observable Markov decision process (POMDP) [12].

The definition of $\mathcal{P}$ implies that the controller has an uncountably infinite number of possible actions. Although this could be circumvented (at the expense of performance) by restricting the powers $P_{n,k,m}^t$ to come from a finite set, the problem would remain very complex due to the continuous-state nature of the SSGs $\gamma_{n,k}^t$. While these SSGs could then be quantized (causing additional performance loss), the problem would still remain computationally intensive, since POMDPs (even with finite states and actions) are PSPACE-complete, i.e., they require both complexity and memory that grow exponentially with the horizon T [23]. To see why, notice from (3) that the solution of the problem at every time t depends on the optimal solution at times up to $t{-}1$. Because both terms on the right side of (3) are dependent on $(\boldsymbol{I}^t, \boldsymbol{P}^t)$, however, the solution of the problem at time t also depends on the solution of the problem at time $t+1$, which in turn depends on the solution of the problem at time $t+2$, and so on. In conclusion, the optimal controller is not practical to implement, even under power/SSG quantization.

Consequently, we will turn our attention to (sub-optimal) *greedy* strategies, i.e., those that do not consider the effect of current actions on future utilities. To better understand their performance relative to that of the optimal POMDP, we derive an upper bound on POMDP performance.

[2] For the $d > 1$ case, the Bellman equation is more complicated, and so we omit it for brevity.

A. The "Causal Global Genie" Upper Bound

Our POMDP-performance upper-bound, which we will refer to as the "causal global genie" (CGG), is based on the presumption of *perfect* error-rate feedback of *all* previous user/subchannel combinations, i.e., $\{\epsilon_{n,k,m}^{\tau}\,\forall n,k,\tau \leqslant t-d\}$. For comparison, the ACK/NAK feedback available to the POMDP is a form of *degraded* error-rate feedback on *previously scheduled* user/subchannel combinations. Since, given knowledge of $\epsilon_{n,k,m}^{\tau}$ and $P_{n,k,m}^{\tau}$ for any rate index m, the SSG $\gamma_{n,k}^{\tau}$ can be obtained by simply inverting the error-rate expression $\epsilon_{n,k,m}^{\tau} = a_m e^{-b_m P_{n,k,m}^{\tau}\gamma_{n,k}^{\tau}}$, our genie-aided bound is based, equivalently, on perfect feedback of all previous SSGs $\{\gamma_{n,k}^{\tau}\,\forall n,k,\tau \leqslant t-d\}$. In the sequel, we use $\boldsymbol{\gamma}^t \in [0,\infty)^{NK}$ to denote the collection of all time-t SSGs $\{\gamma_{n,k}^{t}\,\forall k,n\}$, and we define $\boldsymbol{\gamma}_{-\infty}^{t-d} \triangleq \{\boldsymbol{\gamma}^{-\infty},\dots,\boldsymbol{\gamma}^{t-d}\}$.

We characterize the CGG as "global" since it uses feedback from *all* user/subchannel combinations, not just the previously scheduled ones. Although a tighter bound might result if the (perfect) error-rate feedback was restricted to only previously scheduled user/subchannel pairs, the bounding solution would remain a POMDP with an uncountable number of state-action pairs, making it impractical to evaluate. Evaluating the performance of the CGG, however, is straightforward since—under CGG feedback—optimal scheduling and resource maximization can be performed greedily. To see why, notice that, for any scheduling time $t \geqslant 0$, the CGG scheme allocates resources according to the following mixed-integer optimization problem:

$$(\boldsymbol{I}^{t,\mathsf{cgg}},\boldsymbol{P}^{t,\mathsf{cgg}}) = \arg\max_{(\boldsymbol{I}^t,\boldsymbol{P}^t)\in\mathcal{X}} \sum_{n,k,m} \mathrm{E}\Big\{\tilde{U}_{n,k,m}^{t}\big(I_{n,k,m}^{t},P_{n,k,m}^{t}\big) + \sum_{\tau=t+1}^{T}\tilde{U}_{n,k,m}^{\tau}\big(I_{n,k,m}^{\tau,\mathsf{cgg}},P_{n,k,m}^{\tau,\mathsf{cgg}}\big)\Big|\boldsymbol{I}_{-\infty}^{t-d},\boldsymbol{P}_{-\infty}^{t-d},\boldsymbol{\gamma}_{-\infty}^{t-d}\Big\}. \tag{4}$$

Since the choice of $\big\{(\boldsymbol{I}^{t+1,\mathsf{cgg}},\boldsymbol{P}^{t+1,\mathsf{cgg}}),\dots,(\boldsymbol{I}^{T,\mathsf{cgg}},\boldsymbol{P}^{T,\mathsf{cgg}})\big\}$ does not depend on the choice of $(\boldsymbol{I}^{t,\mathsf{cgg}},\boldsymbol{P}^{t,\mathsf{cgg}})$, the previous optimization problem simplifies to

$$(\boldsymbol{I}^{t,\mathsf{cgg}},\boldsymbol{P}^{t,\mathsf{cgg}}) = \arg\max_{(\boldsymbol{I}^t,\boldsymbol{P}^t)\in\mathcal{X}} \sum_{n,k,m} \mathrm{E}\Big\{\tilde{U}_{n,k,m}^{t}\big(I_{n,k,m}^{t},P_{n,k,m}^{t}\big)\Big|\boldsymbol{\gamma}_{-\infty}^{t-d}\Big\}. \tag{5}$$

In the following lemma, we formally establish that the utility achieved by the CGG upper-bounds that achieved by the optimal POMDP controller with ACK/NAK feedback.

Lemma 1: Given arbitrary past allocations $(\boldsymbol{I}_{-\infty}^{t-d},\boldsymbol{P}_{-\infty}^{t-d})$, and the corresponding ACK/NAKs $\boldsymbol{F}_{-\infty}^{t-d}$, the expected total utility for optimal resource allocation under the latter feedback is no

higher than the expected total utility under CGG feedback, i.e.,

$$\sum_{n,k,m}\sum_{\tau=t}^{T}\mathrm{E}\left\{\tilde{U}^{\tau}_{n,k,m}\big(I^{\tau,\mathsf{opt}}_{n,k,m},P^{\tau,\mathsf{opt}}_{n,k,m}\big)\,\Big|\,\mathbb{F}^{t-d}_{-\infty}\right\}\leqslant\sum_{n,k,m}\sum_{\tau=t}^{T}\mathrm{E}\left\{\tilde{U}^{\tau}_{n,k,m}\big(I^{\tau,\mathsf{cgg}}_{n,k,m},P^{\tau,\mathsf{cgg}}_{n,k,m}\big)\,\Big|\,\mathbb{F}^{t-d}_{-\infty}\right\}. \quad (6)$$

The proof of the above lemma follows the same steps as the proof of [11, Lemma 1], which is omitted here to save space. In the next section, we detail the greedy scheduling and resource allocation problem and propose a near-optimal solution.

IV. Greedy Scheduling and Resource Allocation

The greedy scheduling and resource allocation (GSRA) problem is defined as follows.

$$\begin{aligned}\mathrm{GSRA}\triangleq\max_{\substack{\boldsymbol{I}^t\in\mathcal{I}\\ \boldsymbol{P}^t\in\mathcal{P}}}&\sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{m=1}^{M}I^t_{n,k,m}\,\mathrm{E}\left\{U_{n,k,m}\Big(\big(1-a_m e^{-b_m P^t_{n,k,m}\gamma^t_{n,k}}\big)r_m\Big)\Big|\mathbb{F}^{t-d}_{-\infty}\right\}\\ \text{s.t.}\quad&\sum_{n,k,m}I^t_{n,k,m}\,P^t_{n,k,m}\leqslant X_{\mathsf{con}}.\end{aligned}\quad (7)$$

Note that, in contrast to the T-horizon objective (1), the greedy objective (7) does not consider the effect of $(\boldsymbol{I}^t,\boldsymbol{P}^t)$ on future utility. As stated earlier, we allow $U_{n,k,m}(\cdot)$ to be any real-valued function that is twice differentiable, strictly-increasing, and concave, with $U_{n,k,m}(0)<\infty$. Therefore, $U'_{n,k,m}(\cdot)>0$ and $U''_{n,k,m}(\cdot)\leqslant 0$, using $'$ to denote the derivative.

Since it involves both discrete $(\boldsymbol{I}^t)$ and continuous $(\boldsymbol{P}^t)$ optimization variables, the GSRA problem (7) is a mixed-integer optimization problem. Such problems are generally NP-hard, meaning that polynomial-complexity solutions do not exist. Thus, in Section IV-B, we propose a *near*-optimal algorithm for (7) with polynomial complexity. To better explain that scheme, we first describe, in Section IV-A, a "brute force" optimal solution whose complexity grows exponentially in N, the number of subchannels.

A. Brute-Force Algorithm

The brute-force approach considers all possibilities of $\boldsymbol{I}^t\in\mathcal{I}$, each with the corresponding optimal power allocation. Supposing that $\boldsymbol{I}^t=\boldsymbol{I}$, the optimal power allocation can be found by solving the convex optimization problem

$$\begin{aligned}\max_{\boldsymbol{P}\in\mathcal{P}}&\sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{m=1}^{M}I_{n,k,m}\,\mathrm{E}\left\{U_{n,k,m}\Big(\big(1-a_m e^{-b_m P_{n,k,m}\gamma^t_{n,k}}\big)r_m\Big)\Big|\mathbb{F}^{t-d}_{-\infty}\right\}\\ \text{s.t.}\quad&\sum_{n,k,m}I_{n,k,m}\,P_{n,k,m}\leqslant X_{\mathsf{con}}.\end{aligned}\quad (8)$$

To proceed, we identify the Lagrangian associated with (8) as

$$L^t_{\boldsymbol{I}}(\mu,\boldsymbol{P}) = \Big(\sum_{n,k,m} I_{n,k,m}\,P_{n,k,m} - X_{\textsf{con}}\Big)\mu - \sum_{n,k,m} \mathrm{E}\left\{I_{n,k,m}U_{n,k,m}\big((1-a_m e^{-b_m P_{n,k,m}\gamma^t_{n,k}})r_m\big)\Big|\mathbb{F}^{t-d}_{-\infty}\right\}, \tag{9}$$

which yields the corresponding dual problem

$$\max_{\mu\geqslant 0}\min_{\boldsymbol{P}\in\mathcal{P}} L^t_{\boldsymbol{I}}(\mu,\boldsymbol{P}) = \max_{\mu\geqslant 0} L^t_{\boldsymbol{I}}(\mu,\boldsymbol{P}^*(\mu)) = L^t_{\boldsymbol{I}}(\mu^*_{\boldsymbol{I}},\boldsymbol{P}^*(\mu^*_{\boldsymbol{I}})), \tag{10}$$

where $\mu^*_{\boldsymbol{I}}$ and $\boldsymbol{P}^*(\mu^*_{\boldsymbol{I}})$ denote the optimal Lagrange multiplier and power allocation, respectively.

A detailed solution to (10) is given in [16], and so we describe only the main points here. First, for a given value of the Lagrange multiplier μ, it has been shown that the optimal powers equal

$$P^*_{n,k,m}(\mu) = \begin{cases} \tilde{P}_{n,k,m}(\mu) & \text{if } 0\leqslant\mu\leqslant a_m b_m r_m U'_{n,k,m}\big((1-a_m)r_m\big)\,\mathrm{E}\left\{\gamma^t_{n,k}\middle|\mathbb{F}^{t-d}_{-\infty}\right\} \\ 0 & \text{otherwise,}\end{cases} \tag{11}$$

where $\tilde{P}_{n,k,m}(\mu)$ is defined as the (unique) solution to

$$\mu = a_m b_m r_m\,\mathrm{E}\left\{U'_{n,k,m}\big((1-a_m e^{-b_m\tilde{P}_{n,k,m}(\mu)\gamma^t_{n,k}})r_m\big)\gamma^t_{n,k}e^{-b_m\tilde{P}_{n,k,m}(\mu)\gamma^t_{n,k}}\Big|\mathbb{F}^{t-d}_{-\infty}\right\}. \tag{12}$$

Then, for a given $\boldsymbol{I}$, the optimal value of μ (i.e., $\mu^*_{\boldsymbol{I}}$) obeys $\mu^*_{\boldsymbol{I}}\in[\mu_{\textsf{min}},\mu_{\textsf{max}}]\subset(0,\infty)$, where

$$\mu_{\textsf{min}} = \min_{n,k,m} a_m b_m r_m\,\mathrm{E}\left\{U'_{n,k,m}\big((1-a_m e^{-b_m X_{\textsf{con}}\gamma^t_{n,k}})r_m\big)\gamma^t_{n,k}e^{-b_m X_{\textsf{con}}\gamma^t_{n,k}}\Big|\mathbb{F}^{t-d}_{-\infty}\right\}, \tag{13}$$

$$\mu_{\textsf{max}} = \max_{n,k,m} a_m b_m r_m U'_{n,k,m}\big((1-a_m)r_m\big)\,\mathrm{E}\left\{\gamma^t_{n,k}\middle|\mathbb{F}^{t-d}_{-\infty}\right\}, \tag{14}$$

and satisfies $\sum_{n,k,m} I_{n,k,m}\,P^*_{n,k,m}(\mu^*_{\boldsymbol{I}}) = X_{\textsf{con}}$.

Based on (11)-(14), Table I details the brute-force steps for a given $\boldsymbol{I}$. In the end, for a specified tolerance κ, these steps find $\underline{\mu}$ and $\bar{\mu}$ such that $\mu^*_{\boldsymbol{I}}\in[\underline{\mu},\bar{\mu}]$ and $\bar{\mu}-\underline{\mu}<\kappa$. Using an approximation of $\mu^*_{\boldsymbol{I}}$ that lies in $[\underline{\mu},\bar{\mu}]$, the corresponding utility is guaranteed to be no less than $\kappa X_{\textsf{con}}$ from the optimal (for the given $\boldsymbol{I}$). Therefore, by adjusting κ, one can achieve a performance arbitrarily close to the optimum. Since $|\mathcal{I}| = (KM+1)^N$ values of $\boldsymbol{I}$ must be considered, the total complexity of the brute-force approach—in terms of the number of times (12) must be solved—can be shown to be

$$\left\lceil \log_2\!\left(\tfrac{\mu_{\textsf{max}}-\mu_{\textsf{min}}}{\kappa}\right)\right\rceil \times (KM+1)^{N-1}NKM, \tag{15}$$

which grows exponentially with N.

B. Proposed Algorithm

We propose to attack the mixed-integer GSRA problem (7) using the well known *Lagrangian relaxation* approach [22]. In doing so, we relax the domain of the scheduling variables $I^t_{n,k,m}$ from the set $\{0,1\}$ to the interval $[0,1]$, allowing the application of low-complexity dual optimization techniques. Although the solution to the relaxed problem does not necessarily coincide with that of the original greedy problem (7), we establish in the sequel that the corresponding performance loss is very small, and in some cases zero.

The relaxed version of the greedy problem (7) is

$$\begin{aligned} \text{rGSRA} \triangleq \max_{\substack{\boldsymbol{I}^t\in\mathcal{I}_c \\ \boldsymbol{P}^t\in\mathcal{P}}} &\sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{m=1}^{M} I^t_{n,k,m}\,\mathrm{E}\left\{U_{n,k,m}\Big((1-a_m e^{-b_m P^t_{n,k,m}\gamma_{n,k}})r_m\Big)\Big|\mathbb{F}^{t-d}_{-\infty}\right\} \\ \text{s.t.} \quad &\sum_{n,k,m} I^t_{n,k,m}\,P^t_{n,k,m} \leqslant X_{\mathsf{con}}, \end{aligned} \tag{16}$$

where $\mathcal{I}_c \triangleq \left\{\boldsymbol{I}\in[0,1]^{NKM} : \sum_{k,m} I_{n,k,m} \leqslant 1\ \forall n\right\}$. Although (16) is a non-convex optimization problem due to non-convex constraints, it can be converted into a convex optimization problem by using the new set of variables $(\boldsymbol{I}^t,\boldsymbol{x}^t)$, where $x^t_{n,k,m} \triangleq I^t_{n,k,m}\,P^t_{n,k,m}$. In this case, we have

$$\text{rGSRA} = \min_{\substack{\boldsymbol{x}^t\succeq 0 \\ \boldsymbol{I}^t\in\mathcal{I}_c}} \sum_{n,k,m} I^t_{n,k,m}\,B^t_{n,k,m}(I^t_{n,k,m},x^t_{n,k,m}) \quad \text{s.t.} \sum_{n,k,m} x^t_{n,k,m} \leqslant X_{\mathsf{con}}, \tag{17}$$

where $\boldsymbol{x}^t \in \mathbb{R}^{NKM}$ denotes the collection of all time-t variables $\{x^t_{n,k,m}\}$, $\boldsymbol{x}^t \succeq 0$ denotes element-wise non-negativity, and $B^t_{n,k,m}(\cdot,\cdot)$ is defined as

$$B^t_{n,k,m}(y_1,y_2) \triangleq \begin{cases} -\,\mathrm{E}\left\{U_{n,k,m}\Big((1-a_m e^{-b_m\gamma^t_{n,k}\,y_2/y_1})r_m\Big)\Big|\mathbb{F}^{t-d}_{-\infty}\right\} & \text{if } y_1\neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

The modified problem (17) is a convex optimization problem and can be solved using a dual optimization approach with zero duality gap. In particular, the dual problem can be written as

$$\begin{aligned} &\max_{\mu\geqslant 0}\min_{\substack{\boldsymbol{x}^t\succeq 0 \\ \boldsymbol{I}^t\in\mathcal{I}_c}} L(\mu,\boldsymbol{I}^t,\boldsymbol{x}^t) = \max_{\mu\geqslant 0}\min_{\boldsymbol{I}^t\in\mathcal{I}_c} L(\mu,\boldsymbol{I}^t,\boldsymbol{x}^{t,*}(\mu,\boldsymbol{I}^t)) \\ &= \max_{\mu\geqslant 0} L(\mu,\boldsymbol{I}^{t,*}(\mu),\boldsymbol{x}^{t,*}(\mu,\boldsymbol{I}^{t,*}(\mu))) = L(\mu^*,\boldsymbol{I}^{t,*}(\mu^*),\boldsymbol{x}^{t,*}(\mu^*,\boldsymbol{I}^{t,*}(\mu^*))), \end{aligned} \tag{19}$$

where

$$L(\mu,\boldsymbol{I}^t,\boldsymbol{x}^t) \triangleq \sum_{n,k,m} I^t_{n,k,m}\,B^t_{n,k,m}(I^t_{n,k,m},x^t_{n,k,m}) + \Big(\sum_{n,k,m} x^t_{n,k,m} - X_{\mathsf{con}}\Big)\mu, \tag{20}$$

where $\boldsymbol{x}^*(\mu, \boldsymbol{I})$ is the optimal $\boldsymbol{x}$ for a given $(\mu, \boldsymbol{I})$, where $\boldsymbol{I}^*(\mu)$ denotes the optimal $\boldsymbol{I} \in \mathcal{I}_c$ for a given μ, and where μ^* denotes the optimal $\mu \geqslant 0$.

A detailed solution to this problem was given in [16], and so we describe only the main points here. For given values of μ and $\boldsymbol{I}^t$, we have $x^{t,*}_{n,k,m}(\mu, \boldsymbol{I}^t) = I^t_{n,k,m}\, P^{t,*}_{n,k,m}(\mu)$, where

$$P^{t,*}_{n,k,m}(\mu) = \begin{cases} \tilde{P}^t_{n,k,m}(\mu) & \text{if } 0 \leqslant \mu \leqslant a_m b_m r_m U'_{n,k,m}\big((1-a_m)r_m\big)\, \mathrm{E}\left\{\gamma^t_{n,k} \middle| \mathbb{F}^{t-d}_{-\infty}\right\} \\ 0 & \text{otherwise}, \end{cases} \tag{21}$$

and where $\tilde{P}^t_{n,k,m}(\mu)$ is defined as the (unique) solution to

$$\mu = a_m b_m r_m\, \mathrm{E}\left\{U'_{n,k,m}\Big(\big(1 - a_m e^{-b_m \tilde{P}^t_{n,k,m}(\mu)\gamma^t_{n,k}}\big) r_m\Big) \gamma^t_{n,k} e^{-b_m \tilde{P}^t_{n,k,m}(\mu)\gamma^t_{n,k}} \middle| \mathbb{F}^{t-d}_{-\infty}\right\}. \tag{22}$$

To give equations that govern $\boldsymbol{I}^{t,*}(\mu)$ for a given μ, we first define

$$V^t_{n,k,m}(\mu, P^{t,*}_{n,k,m}(\mu)) \triangleq -\,\mathrm{E}\left\{U_{n,k,m}\Big(\big(1 - a_m e^{-b_m P^{t,*}_{n,k,m}(\mu)\gamma_{n,k}}\big) r_m\Big) \middle| \mathbb{F}^{t-d}_{-\infty}\right\} + \mu P^*_{n,k,m}(\mu) \tag{23}$$

$$S^t_n(\mu) \triangleq \left\{(k,m) = \underset{(k',m')}{\operatorname{argmin}}\, V^t_{n,k',m'}(\mu, P^{t,*}_{n,k',m'}(\mu)) : V^t_{n,k,m}(\mu, P^{t,*}_{n,k,m}(\mu)) \leqslant 0\right\}. \tag{24}$$

If $S^t_n(\mu)$ is a null or a singleton set, then the optimal schedule on subchannel n is given by

$$I^{t,*}_{n,k,m}(\mu) = \begin{cases} 1 & (k,m) \in S^t_n(\mu) \\ 0 & \text{otherwise}. \end{cases} \tag{25}$$

However, if $S^t_n(\mu)$ has cardinality greater than one, then multiple (k,m) combinations can be scheduled simultaneously while achieving the optimal value of the Lagrangian. In particular, if $S^t_n(\mu) = \{(k_1(n), m_1(n)), \ldots, (k_{|S^t_n(\mu)|}(n), m_{|S^t_n(\mu)|}(n))\}$, then

$$I^{t,*}_{n,k,m}(\mu) = \begin{cases} I_{n,k_i(n),m_i(n)} & \text{if } (k,m) = (k_i(n), m_i(n)) \text{ for some } i \in \{1, \ldots, |S^t_n(\mu)|\} \\ 0 & \text{otherwise}, \end{cases} \tag{26}$$

where the vector $\big[I_{n,k_1(n),m_1(n)}, \ldots, I_{n,k_{|S^t_n(\mu)|}(n), m_{|S^t_n(\mu)|}(n)}\big]$ lies anywhere in the unit-$(|S^t_n(\mu)|-1)$ simplex, i.e., it lives within the region $[0,1]^{|S^t_n(\mu)|}$ and satisfies $\sum_{i=1}^{|S^t_n(\mu)|} I_{n,k_i(n),m_i(n)} = 1$. Finally, the optimal Lagrange multiplier μ (i.e., μ^*) is such that $\mu^* \in [\mu_{\mathsf{min}}, \mu_{\mathsf{max}}] \subset (0, \infty)$ and

$$\sum_{n,k,m} I^{t,*}_{n,k,m}(\mu^*)\, P^{t,*}_{n,k,m}(\mu^*) = X_{\mathsf{con}}, \tag{27}$$

where μ_{min} and μ_{max} were given in (13) and (14), respectively.

For several fixed values of μ, the proposed algorithm minimizes the relaxed Lagrangian (20) over $(\boldsymbol{I}^t, \boldsymbol{x}^t)$ (or, equivalently, over $(\boldsymbol{I}^t, \boldsymbol{P}^t)$) to obtain candidate solutions for the original greedy

problem (7). If, for a given μ, $|S_n^t(\mu)| \leqslant 1$ for all n (i.e., the candidate employs at most one user/MCS per subchannel), then the candidate solution is admissible for the non-relaxed problem, and thus retained by the proposed algorithm. If, on the other hand, $|S_n^t(\mu)| > 1$ for some n (i.e., the candidate employs more than one user/MCS on some subchannels), then the proposed algorithm transforms the candidate into an admissible solution as follows:

$$I_{n,k,m}^{t,\text{pro}}(\mu) = \begin{cases} 1 & (k,m) = \operatorname{argmin}_{(k',m') \in S_n^t(\mu)} P_{n,k',m'}^{t,*}(\mu) \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

The following lemma then states an important property of these fixed-μ admissible solutions.

Lemma 2: For any given value of μ, let the power allocation $\boldsymbol{P}^{t,*}(\mu)$ be given by (21), let the user-MCS allocation $\boldsymbol{I}^{t,\text{pro}}(\mu)$ be given by (28), and let the total power allocation be defined as $X_{\text{tot}}^{t,\text{pro}}(\mu) \triangleq \sum_{n,k,m} I_{n,k,m}^{t,\text{pro}}(\mu)\, P_{n,k,m}^{t,*}(\mu)$. Then, $X_{\text{tot}}^{t,\text{pro}}(\mu)$ is monotonically decreasing in μ.

Lemma 2 (see [16] for a proof) implies that the optimal value of the Lagrange multiplier μ (i.e., μ^*) is the one that achieves the power constraint $X_{\text{tot}}^{t,\text{pro}}(\mu) = X_{\text{con}}$. To find this μ^*, the proposed algorithm performs a bisection search over $\mu \in [\mu_{\text{min}}, \mu_{\text{max}}]$ that refines the search interval $[\underline{\mu}, \bar{\mu}]$ until $\bar{\mu} - \underline{\mu} < \kappa$, where κ is a user-defined tolerance. Then, between the two schedules $\boldsymbol{I} \in \{\boldsymbol{I}^{t,\text{pro}}(\underline{\mu}), \boldsymbol{I}^{t,\text{pro}}(\bar{\mu})\}$, it chooses the one that maximizes utility, reminiscent of the brute-force algorithm. Table II summarizes the proposed algorithm.

The complexity of the proposed algorithm—in terms of number of times (22) is solved—is

$$\left\lceil \log_2\!\left(\tfrac{\mu_{\text{max}} - \mu_{\text{min}}}{\kappa}\right) \right\rceil \times N(KM+2), \tag{29}$$

which is significantly less than the brute-force complexity in (15). Although the proposed algorithm is sub-optimal, the difference between the optimal GSRA utility U_{GSRA}^* and that attained by the proposed algorithm $\hat{U}_{\text{GSRA}}(\underline{\mu}, \bar{\mu})$, as $\underline{\mu} \to \bar{\mu}$, can be bounded as follows [16]:

$$U_{\text{GSRA}}^* - \lim_{\underline{\mu} \to \bar{\mu}} \hat{U}_{\text{GSRA}}(\underline{\mu}, \bar{\mu}) \leqslant (\mu^* - \mu_{\text{min}})\big(X_{\text{con}} - X_{\text{tot}}^{t,\text{pro}}(\mu^*)\big) \tag{30}$$

$$\leqslant \begin{cases} 0 & \text{if } |S_n(\mu^*)| \leqslant 1 \ \forall n \\ (\mu_{\text{max}} - \mu_{\text{min}}) X_{\text{con}} & \text{otherwise} \end{cases}. \tag{31}$$

In Section VI, we evaluate (30) by simulation, and show that the performance loss is negligible.

V. Updating the Posterior Distributions from ACK/NAK Feedback

In this section, we propose a recursive procedure to compute the posterior pdfs $p(\gamma_{n,k}^t \,|\, \mathbb{F}_{-\infty}^{t-d})$ required by the proposed greedy algorithm in Table II when the channel is first-order[3] Markov.

Let the time-t user-k channel be described by the discrete-time channel impulse response $\boldsymbol{h}_k^t \triangleq [h_{1,k}^t, \ldots, h_{L,k}^t]^\intercal \in \mathbb{C}^L$, where $(\cdot)^\intercal$ denotes transpose. The corresponding frequency-domain subchannel gains $\boldsymbol{H}_k^t \triangleq [H_{1,k}^t, \ldots, H_{N,k}^t]^\intercal \in \mathbb{C}^N$ are then given by

$$\boldsymbol{H}_k^t = \boldsymbol{G}\boldsymbol{h}_k^t, \tag{32}$$

where the OFDMA modulation matrix $\boldsymbol{G} \in \mathbb{C}^{N\times L}$ contains the first L columns of the N-DFT matrix. Assuming additive white Gaussian noise with unit variance, the SSG of subchannel n for user k is given by $\gamma_{n,k}^t = |H_{n,k}^t|^2$, and so we can write

$$p(\gamma_{n,k}^t \,|\, \mathbb{F}_{-\infty}^{t-d}) = \int_{\boldsymbol{h}_k^t} p(\gamma_{n,k}^t \,|\, \boldsymbol{h}_k^t) p(\boldsymbol{h}_k^t \,|\, \mathbb{F}_{-\infty}^{t-d}) \tag{33}$$

with $p(\gamma_{n,k}^t \,|\, \boldsymbol{h}_k^t) = \delta(\gamma_{n,k}^t - |\boldsymbol{e}_n^\intercal \boldsymbol{G}\boldsymbol{h}_k^t|^2)$, where $\delta(\cdot)$ is the Dirac delta and $\boldsymbol{e}_n$ is the n^{th} column of the identity matrix. Using the channel's Markov property and Bayes rule, we find that

$$p(\boldsymbol{h}_k^t \,|\, \mathbb{F}_{-\infty}^{t-d}) = \int_{\boldsymbol{h}_k^{t-d}} p(\boldsymbol{h}_k^t \,|\, \boldsymbol{h}_k^{t-d})\, p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d}) \tag{34}$$

$$p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d}) = \frac{p(\boldsymbol{f}_k^{t-d} \,|\, \boldsymbol{h}_k^{t-d}, \mathbb{F}_{-\infty}^{t-d} \setminus \boldsymbol{f}_k^{t-d})\, p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d} \setminus \boldsymbol{f}_k^{t-d})}{\int_{\bar{\boldsymbol{h}}_k^{t-d}} p(\boldsymbol{f}_k^{t-d} \,|\, \bar{\boldsymbol{h}}_k^{t-d}, \mathbb{F}_{-\infty}^{t-d} \setminus \boldsymbol{f}_k^{t-d})\, p(\bar{\boldsymbol{h}}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d} \setminus \boldsymbol{f}_k^{t-d})}, \tag{35}$$

where $\setminus$ denotes the set-difference operator. Using the fact that $p(\boldsymbol{f}_k^{t-d} \,|\, \boldsymbol{h}_k^{t-d}, \mathbb{F}_{-\infty}^{t-d} \setminus \boldsymbol{f}_k^{t-d}) = p(\boldsymbol{f}_k^{t-d} \,|\, \boldsymbol{h}_k^{t-d}, \boldsymbol{I}^{t-d}, \boldsymbol{p}^{t-d})$, along with the fact that $(\boldsymbol{I}^{t-d}, \boldsymbol{p}^{t-d})$ is a deterministic function of $\mathbb{F}_{-\infty}^{t-2d}$ (and therefore of $\mathbb{F}_{-\infty}^{t-d-1}$), we then have from (35) that

$$p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d}) = \frac{p(\boldsymbol{f}_k^{t-d} \,|\, \boldsymbol{h}_k^{t-d}, \boldsymbol{I}^{t-d}, \boldsymbol{p}^{t-d})\, p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d-1})}{\int_{\bar{\boldsymbol{h}}_k^{t-d}} p(\boldsymbol{f}_k^{t-d} \,|\, \bar{\boldsymbol{h}}_k^{t-d}, \boldsymbol{I}^{t-d}, \boldsymbol{p}^{t-d})\, p(\bar{\boldsymbol{h}}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d-1})}. \tag{36}$$

Using the Markov property again, we get

$$p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d-1}) = \int_{\boldsymbol{h}_k^{t-d-1}} p(\boldsymbol{h}_k^{t-d} \,|\, \boldsymbol{h}_k^{t-d-1})\, p(\boldsymbol{h}_k^{t-d-1} \,|\, \mathbb{F}_{-\infty}^{t-d-1}). \tag{37}$$

Recall that $f_{n,k}^t$, the feedback received about user k on channel n at time t, takes values from the set $\{0, 1, \emptyset\}$, where 0 denotes a NAK, 1 denotes an ACK, and $\emptyset$ denotes that user k was not

[3] The extension to higher-order Markov channels is straightforward.

scheduled on subchannel n at time t. Assuming that, conditioned on $\boldsymbol{h}_k^t$, the feedbacks generated by user k are independent across subchannels, we have

$$p(\boldsymbol{f}_k^t|\boldsymbol{h}_k^t, \boldsymbol{I}^t, \boldsymbol{p}^t) = \prod_{n=1}^{N} p(f_{n,k}^t|\boldsymbol{h}_k^t, \boldsymbol{I}^t, \boldsymbol{p}^t), \tag{38}$$

$$p(f_{n,k}^t = f \,|\, \boldsymbol{h}_k^t, \boldsymbol{I}^t, \boldsymbol{p}^t) = \begin{cases} \sum_m I_{n,k,m}^t a_m e^{-b_m p_{n,k,m}^t \gamma_{n,k}^t} & \text{if } f = 0 \\ \sum_m I_{n,k,m}^t \left(1 - a_m e^{-b_m p_{n,k,m}^t \gamma_{n,k}^t}\right) & \text{if } f = 1 \\ 1 - \sum_m I_{n,k,m}^t & \text{if } f = \emptyset, \end{cases} \tag{39}$$

where $\gamma_{n,k}^t = |H_{n,k}^t|^2$ can be determined from $\boldsymbol{h}_k^t$ via (32). Together, (33)-(39) suggest a method of *recursively* updating the channel distributions, using the new feedback obtained at each time t, which is given in Table III.

We now propose the use of particle filtering [24] to circumvent the evaluation of multidimensional integrals in the recursion of Table III. Particle filtering is a well-known technique that approximates the pdf of a random variable using a suitably chosen probability mass function (pmf). In the sequel, for simplicity of illustrations, we assume a Gauss-Markov model of the form

$$h_{l,k}^{t+1} \;=\; (1-\alpha) h_{l,k}^t + \alpha w_{l,k}^t, \tag{40}$$

where $w_{l,k}^t$ is unit-variance circular Gaussian and $\alpha \in (0, 1]$ is a known constant that determines the fading rate. Here, $w_{l,k}^t$ is assumed to be i.i.d. for all t, l, k. At each time-step t, for $k \in \{1, \ldots K\}$, we use S particles in the approximations

$$\begin{aligned} p(\boldsymbol{h}_k^t \,|\, \mathbb{F}_{-\infty}^{t-d}) &\approx \sum_{i=1}^{S} \nu_k^{t\,|\,t-d}[i]\, \delta(\boldsymbol{h}_k^t - \boldsymbol{h}_k^t[i]), \text{ and} \\ p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d}) &\approx \sum_{i=1}^{S} \nu_k^{t-d\,|\,t-d}[i]\, \delta(\boldsymbol{h}_k^{t-d} - \boldsymbol{h}_k^{t-d}[i]), \end{aligned} \tag{41}$$

where $\boldsymbol{h}_k^t[i] = \left[h_{1,k}^t[i], \ldots, h_{L,k}^t[i]\right]^\mathsf{T} \in \mathbb{C}^L$ denotes the i^{th} (vector) particle, for $i \in \{1, \ldots, S\}$, and $\nu_k^{t_1|t_2}[i] \in \mathbb{R}^+$ is the probability mass assigned to the particle $\boldsymbol{h}_k^{t_1}[i]$ based on the observations received up to time t_2. The steps to recursively compute these particles and their corresponding weights are detailed in Table IV.

Using the approximation in (41), we note that the expectation of any function of subchannel-gain, $\boldsymbol{h}_k^t$, can be found using

$$\mathrm{E}\{A(\boldsymbol{h}_k^t)\} \approx \sum_i \nu_k^{t\,|\,t-d}[i] A\big(\boldsymbol{h}_k^t[i]\big), \tag{42}$$

where $A(\cdot)$ is an arbitrary function. Recalling that the SSG $\gamma_{n,k}^t$ is a deterministic function of the subchannel-gain, $\boldsymbol{h}_k^t$, any function of $\gamma_{n,k}^t$ is also a function of $\boldsymbol{h}_k^t$.

VI. Numerical Results

In this section, we numerically evaluate the performance of the proposed greedy scheduling and resource allocation from Section IV with the posterior update from Section V. For this, we consider an OFDMA system with independent first-order Gauss-Markov channels (40). We assumed, if not otherwise stated, $K = 8$ available users, $N = 32$ OFDMA subchannels, channel fading parameter $\alpha = 10^{-3}$ and impulse response length $L = 2$. We used the modulation matrix $\boldsymbol{G} = \sqrt{\beta}\boldsymbol{F} \in \mathbb{C}^{N\times L}$ (recall (32)), where $\boldsymbol{F}$ contains the first $L(\leqslant N)$ columns of the unitary N-DFT matrix and $\beta = \frac{N}{L}\frac{2-\alpha}{\alpha}$ ensures that the variance of $H_{n,k}^t$ is unity for all (n,k). Thus, the mean of the SSG $\gamma_{n,k}^t$ was also unity for all (n,k). Since the subchannel-averaged total transmit power equals $\frac{1}{N}\sum_{n,k,m} I_{n,k,m}^t P_{n,k,m}^t = \frac{1}{N}\sum_{n,k,m} X_{n,k,m}^t = X_{\textsf{con}}/N$, it is readily seen that the average per-subchannel signal-to-noise ratio is $\textsf{SNR} \triangleq \mathrm{E}\{\frac{1}{N}\sum_{n,k,m} X_{n,k,m}^t \gamma_{n,k}^t\} = \frac{1}{N}\sum_{n,k,m} X_{n,k,m}^t \mathrm{E}\{\gamma_{n,k}^t\} = X_{\textsf{con}}/N$. For the plots, we averaged 500 realizations, each with 100 time-slots. Of these 100 time-slots, the first 50 were ignored to avoid transient effects.

For illustrative purposes, we assumed uncoded 2^{m+1}-QAM signaling with MCS index $m \in \{1,\ldots,15\}$. In this case, we have $r_m = m+1$ bits per symbol, one symbol per "codeword," and one codeword per packet. In the packet error-rate model $\epsilon = a_m e^{-b_m P\gamma}$, we assumed $a_m = 1$ and $b_m = 1.5/(2^{m+1}-1)$ because the symbol error-rate of a 2^{m+1}-QAM system is well approximated by $\exp(-1.5P\gamma/(2^{m+1}-1))$ in the high-$(P\gamma)$ regime [25] and is ≈ 1 when $P\gamma = 0$. Throughout, we used the identity utility (i.e., $U_{n,k,m}(x) = x$ for all n,k,m) so that the objective was maximization of sum goodput, and we assumed a feedback delay of $d = 1$.

The performance of the proposed greedy algorithm was compared to three reference schemes: fixed-power random user scheduling (FP-RUS), the "causal global genie" (CGG), and the "non-causal global genie" (NCGG). The FP-RUS scheme schedules users uniformly at random, allocates power uniformly across subchannels, and selects the MCS to maximize expected goodput.

The FP-RUS, which makes no use of feedback, should perform no better than any feedback-based scheme. The CGG (recall Section III-A) performs optimal scheduling and resource allocation under perfect knowledge of all SSGs at the previous time-instant (since $d = 1$), i.e., given $\{\gamma_{n,k}^{t-1} \; \forall n, k\}$ at time t. From Lemma 1, we know that the CGG upper-bounds the POMDP. The NCGG is similar to the CGG, but assumes perfect knowledge of all SSGs at *all* times, i.e., given $\{\gamma_{n,k}^{\tau} \; \forall n, k, \tau\}$ at time t. Thus, it provides an upper bound on the CGG that is invariant to fading rate α. The NCGG has a greedy implementation, like the CGG, but without the conditional expectation in (5).

Figure 1 shows a typical realization of instantaneous sum-goodput versus time t, when $\alpha = 10^{-3}$. There, one can see a large gap between the FP-RUS and the CGG, and a much smaller gap between the CGG and the NCGG. The proposed scheme starts without CSI, and initially performs no better than the FP-RUS. From ACK/NAK feedbacks, however, it quickly learns the CSI well enough to perform scheduling and resource allocation at a level that yields sum-goodput much closer to the CGG than to the FP-RUS.

Figure 2 plots average sum-goodput versus the number of particles S used to update the posterior distributions in the proposed greedy scheme (recall Section V). There we see that the performance of the proposed scheme increases with S, but shows little improvement for $S > 30$. Thus, $S = 30$ particles were used to construct the other plots. Remarkably, with only $S = 5$ particles, the proposed algorithm captures a significant portion of the maximum possible goodput gain over the FP-RUS.

Figure 3 plots average sum-goodput versus the fading rate α. There we see that, at low fading rates (i.e., small α), the proposed greedy scheme achieves an average sum-goodput that is much higher than the FP-RUS and, in fact, not far from the CGG upper bound. For instance, at $\alpha = 10^{-4}$, the sum-goodput attained by the proposed scheme is 92% of the upper bound and 170% of that attained by the FP-RUS. As the fading rate α increases, we see that the sum-goodput attained by the proposed scheme decreases, and eventually converges to that of the FP-RUS. This behavior is due to the fact that, as α increases, it becomes more difficult to predict the SSGs using delayed ACK/NAK feedback, thereby compromising the scheduling-and-resource-allocation decisions that are made based on the predicted SSGs. In fact, one can even observe a gap between the CGG and NCGG for large α because, even with delayed perfect-SSG feedback, the current SSGs are difficult to predict.

Figure 3 reveals a gap between the proposed scheme and the CGG bound that persists as $\alpha \to 0$. This non-vanishing gap can be attributed—at least in part—to *greedy* scheduling under ACK/NAK feedback. Intuitively, we have the following explanation. Because the inferred SSG-distributions of not-recently-scheduled users quickly revert to their apriori form, the proposed greedy algorithm will continue to schedule users as long as their SSGs remain better than the apriori value. There may exist, however, not-recently-scheduled users with far better SSGs who remain invisible to the proposed scheme, only because they have not recently been scheduled.

Figures 4 and 5 plot average sum-goodput versus the number of subchannels (i.e., total bandwidth) N. In Fig. 4, the total BS power X_{con} is scaled with N such that the per-subchannel SNR remains fixed at 10dB, whereas, in Fig. 5, the total BS power X_{con} remains invariant to the bandwidth N, and is set such that per-subchannel $\mathsf{SNR} = 10$dB for $N = 32$. In both cases, the average sum-goodput increases with bandwidth N, as expected, since the availability of more subchannels increases not only scheduling flexibility, but also the possibility of stronger subchannels, which can be exploited by the BS. In Fig. 4, where the per-subchannel SNR is fixed, the sum-goodput increases linearly with bandwidth N, as expected. In all cases, the proposed greedy scheme captures about 80% of the sum-goodput gain achievable over the FP-RUS.

Figure 6 plots average sum-goodput versus the number of available users K. It shows that, as K increases, the average sum-goodputs achieved by the NCGG, CGG, and the proposed greedy schemes increase, whereas that achieved by the FP-RUS remains constant. This behavior results because, with the former schemes, the availability of more users can be exploited to schedule users with stronger subchannels, whereas with the FP-RUS scheme, this advantage is lost due to the complete lack of information about the users' instantaneous channel conditions. Figure 6 also suggests that, as K increases, the sum-goodput of the proposed greedy scheme saturates. This can be attributed to the fact that the proposed greedy algorithm can only track the channels of recently scheduled users, and thus cannot benefit directly from the growing pool of not-recently-scheduled users.

In Figure 7, the top subplot shows average sum-goodput versus SNR, while the bottom subplot shows the average value of the bound (30) on the optimality gap of our proposed approach to the GSRA problem, also versus SNR. The top plot shows that, as the SNR increases, the proposed greedy scheme continues to perform much closer to the NCGG/CGG bounds than it does to the FP-RUS scheme. The bottom plot establishes that the sum-goodput loss due to the sub-optimality

in the algorithm used to attack the GSRA problem is negligible, e.g., at most 0.0025% over all SNR.

VII. Conclusion

In this paper, we considered the problem of joint scheduling and resource allocation in the OFDMA downlink under ACK/NAK feedback, with the goal of maximizing an expected long-term goodput-based utility subject to an instantaneous sum-power constraint. First, we established that the optimal solution to the problem is a partially observable Markov decision process (POMDP), which is impractical to implement. Consequently, we proposed a greedy approach to joint scheduling and resource allocation based on the posterior distributions of the squared sub-channel gain (SSG) for every user/subchannel pair, which has polynomial complexity. Next, for Markov channels, we outlined a recursive method to update the posterior SSG distributions from the ACK/NAK feedbacks received at each time-slot, and proposed an efficient implementation based on particle filtering. To gauge the performance of our greedy scheme relative to that of the optimal POMDP (which is impossible to implement), we derived a performance upper-bound on POMDP, known as the causal global genie (CGG). Numerical experiments suggest that our greedy scheme achieves a significant fraction of the maximum possible performance gain over fixed-power random user scheduling (FP-RUS), despite its low-complexity implementation. For example, a representative simulation using $N = 32$ OFDMA subchannels, $K = 8$ available users, SNR$= 10$dB, and $S = 30$ particles, shows that the sum-goodput of the proposed scheme is 92% of the upper bound and 170% of that attained by the FP-RUS (see Fig. 3).

References

[1] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks—Parts I and II," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 614–634, Mar. 2005.

[2] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit and power allocation," *IEEE J. Select. Areas In Commun.*, vol. 17, pp. 1747–1758, Oct. 1999.

[3] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Select. Areas In Commun.*, vol. 21, pp. 171–178, Feb. 2003.

[4] T. J. Willink and P. H. Wittke, "Optimization and performance evaluation of multicarrier transmission," *IEEE Trans. Inform. Theory*, vol. 43, pp. 426–440, Mar. 1997.

[5] L. M. C. Hoo, B. Halder, J. Tellado, and J. M. Cioffi, "Multiuser transmit optimization for multicarrier broadcast channels: Asymptotic FDMA capacity region and algorithms," *IEEE Trans. Commun.*, vol. 52, pp. 922–930, Jun. 2004.

[6] I. C. Wong and B. L. Evans, “Optimal OFDMA resource allocation with linear complexity to maximize ergodic weighted sum capacity,” *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, pp. III–601 –III–604, 15-20 Apr. 2007.

[7] K. Seong, M. Mohseni, and J. M. Cioffi, “Optimal resource allocation for OFDMA downlink systems,” *Proc. IEEE Int. Symposium Inform. Theory*, pp. 1394–1398, Jul. 2006.

[8] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, “An overview of limited feedback in wireless communications systems,” *IEEE J. Select. Areas In Commun.*, vol. 26, pp. 1341–1365, Oct. 2008.

[9] D. P. Bertsekas and R. G. Gallager, *Data Networks*. Prentice Hall, 2nd ed., 1992.

[10] A. K. Karmokar, D. V. Djonin, and V. K. Bhargava, “POMDP-based coding rate adaptation for type-I hybrid ARQ systems over fading channels with memory,” *IEEE Trans. Wireless Commun.*, vol. 5, pp. 3512–3523, Dec. 2006.

[11] R. Aggarwal, P. Schniter, and C. E. Koksal, “Rate adaptation via link-layer feedback for goodput maximization over a time-varying channel,” *IEEE Trans. Wireless Commun.*, vol. 8, pp. 4276–4285, Aug. 2009.

[12] G. E. Monahan, “A survey of partially observable Markov decision processes: Theory, models, and algorithms,” *Management Science*, vol. 28, pp. 1–16, Jan. 1982.

[13] M. Falkner, M. Devetsikiotis, and I. Lambadaris, “An overview of pricing concepts for broadband IP networks,” *IEEE Commun. Surveys Tutorials*, vol. 3, pp. 2–13, second quarter 2000.

[14] L. A. DaSilva, “Pricing for QoS-enabled networks: A survey,” *IEEE Commun. Surveys Tutorials*, vol. 3, pp. 2–8, second quarter 2000.

[15] S. Shenker, “Fundamental design issues for the future internet,” *IEEE J. Select. Areas In Commun.*, vol. 13, pp. 1176–1188, Sept. 1995.

[16] R. Aggarwal, M. Assaad, C. E. Koksal, and P. Schniter, “Joint scheduling and resource allocation in the OFDMA downlink: Utility maximization under imperfect channel-state information,” *IEEE Trans. Signal Processing*, vol. 59, pp. 5589–5604, Nov. 2011.

[17] M. A. Haleem and R. Chandramouli, “Adaptive downlink scheduling and rate selection: a cross-layer design,” *IEEE J. Select. Areas In Commun.*, vol. 23, pp. 1287–1297, Jun. 2005.

[18] R. Wang and V. K. N. Lau, “Robust optimal cross-layer designs for TDD-OFDMA systems with imperfect CSIT and unknown interference – state-space approach based on 1-bit ACKNAK feedbacks,” *IEEE Trans. Commun.*, vol. 56, pp. 754–761, May 2008.

[19] Z. K. M. Ho, V. K. N. Lau, and R. S.-K. Cheng, “Cross-layer design of FDD-OFDM systems based on ACK/NAK feedbacks,” *IEEE Trans. Inform. Theory*, vol. 55, pp. 4568–4584, Oct. 2009.

[20] I. C. Wong and B. L. Evans, “Optimal resource allocation in the OFDMA downlink with imperfect channel knowledge,” *IEEE Trans. Commun.*, vol. 57, pp. 232–241, Jan. 2009.

[21] Y. J. Zhang and S. C. Liew, “Proportional fairness in multi-channel multi-rate wireless networks–Part II: The case of time-varying channels with application to OFDM systems,” *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3457–3467, Sept. 2008.

[22] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982.

[23] C. H. Papadimitriou and J. N. Tsitsiklis, “The complexity of Markov decision processes,” *Math. Operations Res.*, vol. 12, pp. 441–450, Aug. 1987.

[24] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking,” *IEEE Trans. Signal Processing*, vol. 50, pp. 174–188, Feb. 2002.

[25] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 5th ed., 2008.

TABLE I
BRUTE-FORCE STEPS FOR A GIVEN $\boldsymbol{I}$

1) Initialize $\underline{\mu} = \mu_{\text{min}}$ and $\bar{\mu} = \mu_{\text{max}}$.
2) Set $\mu = \frac{\underline{\mu}+\bar{\mu}}{2}$.
3) For each (n, k, m),
    a) Use (11)-(12) to obtain $P^*_{n,k,m}(\mu)$.
4) Calculate $X^*_{\text{tot}}(\boldsymbol{I}, \mu) \triangleq \sum_{n,k,m} I_{n,k,m}\, P^*_{n,k,m}(\mu)$.
5) If $X^*_{\text{tot}}(\boldsymbol{I}, \mu) > X_{\text{con}}$, set $\underline{\mu} = \mu$, otherwise set $\bar{\mu} = \mu$.
6) If $\bar{\mu} - \underline{\mu} > \kappa$, go to step 2), else proceed to step 7).
7) If $X^*_{\text{tot}}(\boldsymbol{I}, \bar{\mu}) \neq X^*_{\text{tot}}(\boldsymbol{I}, \underline{\mu})$, set $\lambda = \frac{X^*_{\text{tot}}(\boldsymbol{I},\underline{\mu}) - X_{\text{con}}}{X^*_{\text{tot}}(\boldsymbol{I},\underline{\mu}) - X^*_{\text{tot}}(\boldsymbol{I},\bar{\mu})}$, otherwise set $\lambda = 0$.
8) Set $\hat{\mu}_{\boldsymbol{I}} = \bar{\mu}$. The best power allocation is given by $\hat{\boldsymbol{P}}(\boldsymbol{I}) = \lambda \boldsymbol{P}^*(\bar{\mu}) + (1-\lambda)\boldsymbol{P}^*(\underline{\mu})$, and $\hat{L}_{\boldsymbol{I}} = L^t_{\boldsymbol{I}}(\bar{\mu}, \hat{\boldsymbol{P}}(\boldsymbol{I}))$ gives the best Lagrangian value.

TABLE II
PROPOSED GREEDY ALGORITHM

1) Initialize $\underline{\mu} = \mu_{\text{min}}$ and $\bar{\mu} = \mu_{\text{max}}$.
2) Set $\mu = \frac{\underline{\mu}+\bar{\mu}}{2}$.
3) For each subchannel $n = 1, \ldots, N$:
    a) For each (k, m),
        i) Use (21)-(22) to calculate $P^{t,*}_{n,k,m}(\mu)$.
        ii) Use $P^{t,*}_{n,k,m}(\mu)$ to calculate $V^t_{n,k,m}(\mu, P^{t,*}_{n,k,m}(\mu))$ via (23).
    b) Calculate $S^t_n(\mu)$ using (24).
4) Find $\boldsymbol{I}^{t,\text{pro}}(\mu)$ using (28).
5) Calculate $X^{t,\text{pro}}_{\text{tot}}(\mu) = \sum_{n,k,m} I^{t,\text{pro}}_{n,k,m}(\mu)\, P^{t,*}_{n,k,m}(\mu)$.
6) If $X^{t,\text{pro}}_{\text{tot}}(\mu) > X_{\text{con}}$, set $\underline{\mu} = \mu$, otherwise set $\bar{\mu} = \mu$.
7) If $\bar{\mu} - \underline{\mu} > \kappa$, go to step 2), else proceed to step 8).
8) Now we have $\mu^* \in [\underline{\mu}, \bar{\mu}]$ and $\bar{\mu} - \underline{\mu} < \kappa$. For both $\boldsymbol{I} = \boldsymbol{I}^{t,\text{pro}}(\underline{\mu})$ and $\boldsymbol{I} = \boldsymbol{I}^{t,\text{pro}}(\bar{\mu})$ (since they may differ), calculate $\hat{\boldsymbol{P}}(\boldsymbol{I})$ and $\hat{L}_{\boldsymbol{I}}$ as described for the brute force algorithm.
9) Choose $\hat{\boldsymbol{I}}^t = \operatorname{argmin}_{\boldsymbol{I} \in \{\boldsymbol{I}^{t,\text{pro}}(\underline{\mu}),\, \boldsymbol{I}^{t,\text{pro}}(\bar{\mu})\}} \hat{L}_{\boldsymbol{I}}$ as the user-MCS allocation and $\hat{\boldsymbol{P}}^t = \hat{\boldsymbol{P}}(\hat{\boldsymbol{I}}^t)$ as the associated power allocation.

TABLE III
RECURSIVE UPDATE OF CHANNEL POSTERIORS

At time t, for each user k, the pdf $p(\boldsymbol{h}_k^{t-d-1} \,|\, \mathbb{F}_{-\infty}^{t-d-1})$ is available from the previous time-instant. The user-k recursion is then

1) Observe new feedbacks $\boldsymbol{f}_k^{t-d} \in \{0, 1, \emptyset\}^N$.
2) Compute $p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d-1})$ using (37).
3) Compute $p(\boldsymbol{f}_k^{t-d} \,|\, \boldsymbol{h}_k^{t-d}, \boldsymbol{I}^{t-d}, \boldsymbol{P}^{t-d})$ using the error-rate rule (38)-(39).
4) Using the distributions obtained in steps 2) and 3), compute $p(\boldsymbol{h}_k^{t-d} \,|\, \mathbb{F}_{-\infty}^{t-d})$ via Bayes-rule step in (36).
5) Compute $p(\boldsymbol{h}_k^{t} \,|\, \mathbb{F}_{-\infty}^{t-d})$ using the Markov-prediction step (34).
6) For each n, compute $p(\gamma_{n,k}^{t} \,|\, \mathbb{F}_{-\infty}^{t-d})$ via (33).

TABLE IV
PARTICLE FILTERING STEPS

Let the system begin at time-instant t_0. If $t \in \{t_0, \ldots, t_0 + d - 1\}$:

1) Initialize $\{h_{l,k}^t[i] \ \forall i, k, l\}$ by drawing i.i.d. samples from $\mathcal{CN}(0, \frac{\alpha}{2-\alpha})$.
2) Set the importance weights $\nu_k^{t|t}[i] = \frac{1}{S} \ \forall k, i$.

For any other time-instant t ($\geqslant t_0 + d$):

1) Using the previous samples $\{h_{l,k}^{\tau}[i] \ \forall l, k, i, \tau : \tau \leqslant t - d\}$, obtain new samples according to the underlying Markov model as follows
$$h_{l,k}^t[i] = (1-\alpha)^d \, h_{l,k}^{t-d}[i] + \alpha \sum_{j=0}^{d-1} (1-\alpha)^j \, y_{l,k}^{t-j}[i], \ \forall i, l, k,$$
where $y_{l,k}^{t-j}[i]$ is drawn i.i.d from $\mathcal{CN}(0, 1)$ for all i, l, k, j.
2) For each user k,
   a) Using the received feedbacks $\boldsymbol{f}_k^{t-d}$ and the set of importance weights from time $(t - d - 1)$, i.e., $\{\nu_k^{t-d-1|t-d-1}[i] \ \forall i\}$, compute the new set of importance weights at time $(t - d)$ using
   $$\nu_k^{t-d|t-d}[i] = \nu_k^{t-d-1|t-d-1}[i] \times p\big(\boldsymbol{f}_k^{t-d} \,\big|\, \boldsymbol{h}_k^{t-d} = \boldsymbol{h}_k^{t-d}[i], \boldsymbol{I}^{t-d}, \boldsymbol{P}^{t-d}\big),$$
   for all i, where $p\big(\boldsymbol{f}_k^{t-d} \,\big|\, \boldsymbol{h}_k^{t-d} = \boldsymbol{h}_k^{t-d}[i], \boldsymbol{I}^{t-d}, \boldsymbol{P}^{t-d}\big)$ is given by (38)-(39).
   b) Normalize the weights via
   $$\nu_k^{t-d|t-d}[i] \leftarrow \frac{\nu_k^{t-d|t-d}[i]}{\sum_j \nu_k^{t-d|t-d}[j]} \ \forall i.$$
   c) Compute the weights for the posterior distribution, $p(\boldsymbol{h}_k^t \,|\, \mathbb{F}_{-\infty}^{t-d})$ using
   $$\nu_k^{t|t-d}[i] = \sum_{j=1}^{S} \nu_k^{t-d|t-d}[j] \; p(\boldsymbol{h}_k^t = \boldsymbol{h}_k^t[i] \,|\, \boldsymbol{h}_k^{t-d} = \boldsymbol{h}_k^{t-d}[j]).$$
   d) Normalize the weights via
   $$\nu_k^{t|t-d}[i] \leftarrow \frac{\nu_k^{t|t-d}[i]}{\sum_j \nu_k^{t|t-d}[j]} \ \forall i.$$

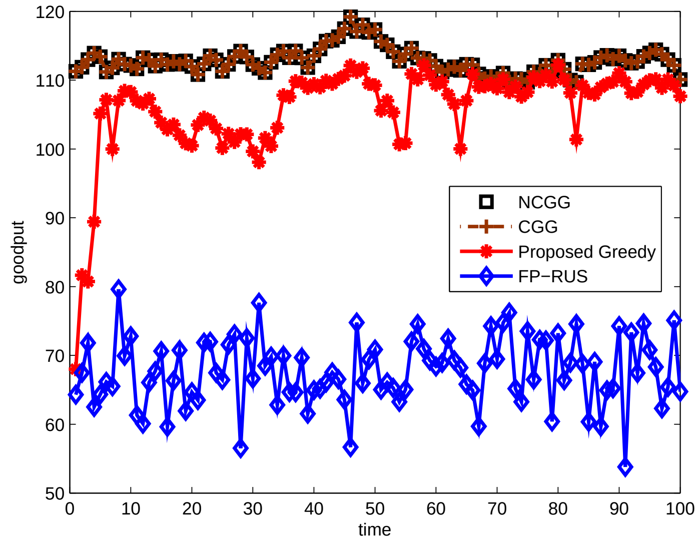

Fig. 1. Typical instantaneous sum-goodput versus time t. Here, $N = 32$, $K = 8$, $\mathsf{SNR} = 10\text{dB}$, $\alpha = 10^{-3}$, and $S = 30$.

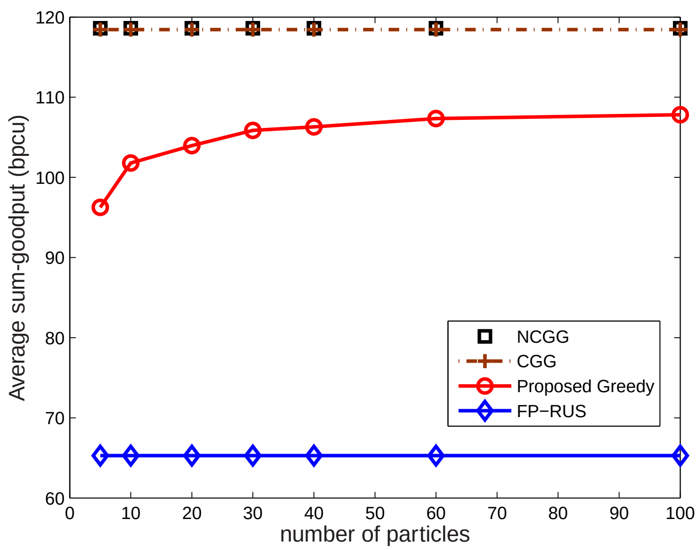

Fig. 2. Average sum-goodput versus the number of particles used to update the channel posteriors. Here, $N = 32$, $K = 8$, $\mathsf{SNR} = 10\text{dB}$, and $\alpha = 10^{-3}$.

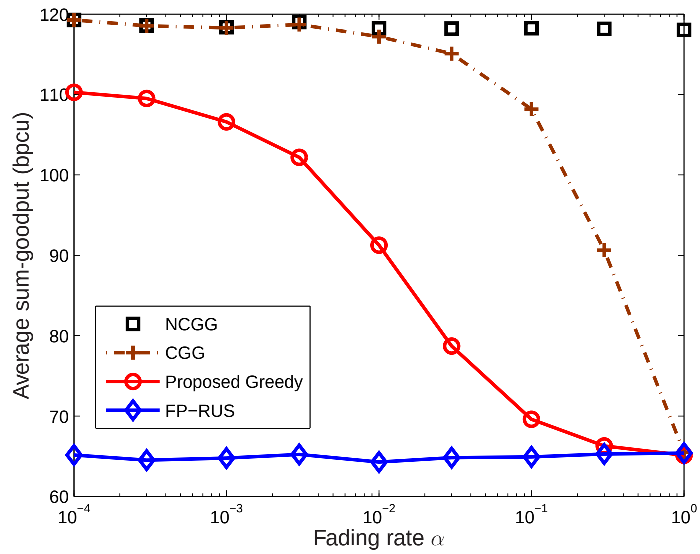

Fig. 3. Average sum-goodput versus fading rate α. Here, $N = 32$, $K = 8$, SNR $= 10$dB, and $S = 30$.

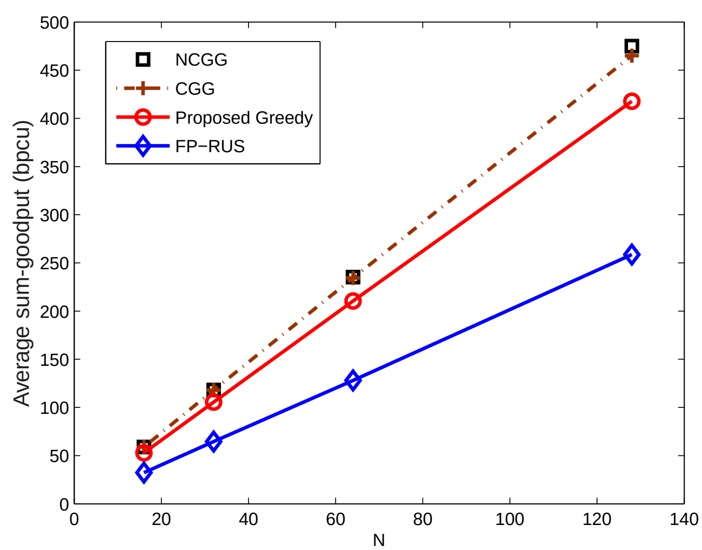

Fig. 4. Average sum-goodput versus number of subchannels N. Here, $K = 8$, SNR $= 10$dB, $\alpha = 10^{-3}$, and $S = 30$.

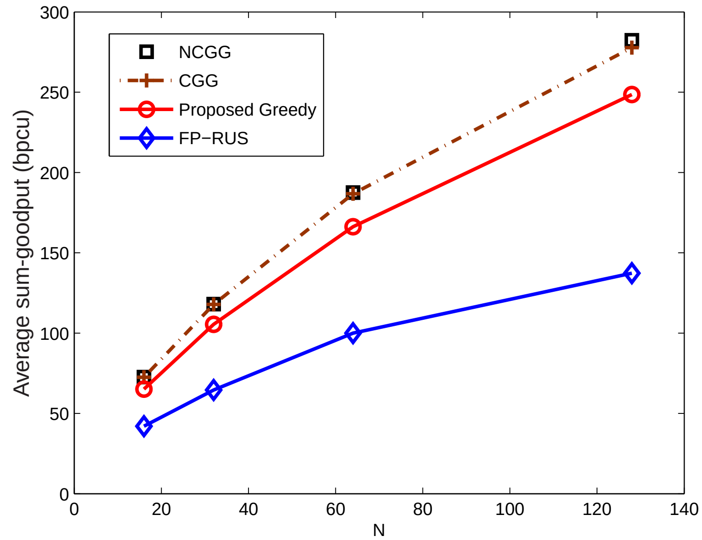

Fig. 5. Average sum-goodput versus number of subchannels N. Here, $K = 8$, X_{con} does not scale with N and it is chosen such that $\mathsf{SNR} = 10\text{dB}$ for $N = 32$, $\alpha = 10^{-3}$, and $S = 30$.

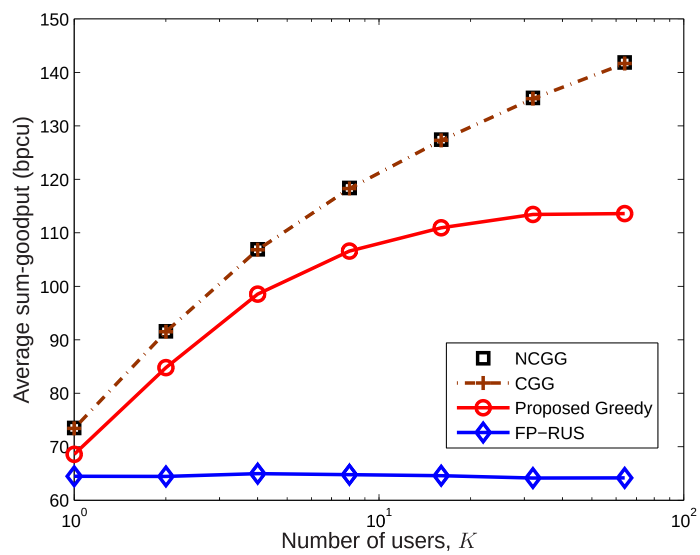

Fig. 6. Average sum-goodput versus number of users. In this plot, $N = 32$, $\mathsf{SNR} = 10\text{dB}$, $\alpha = 10^{-3}$, and $S = 30$.

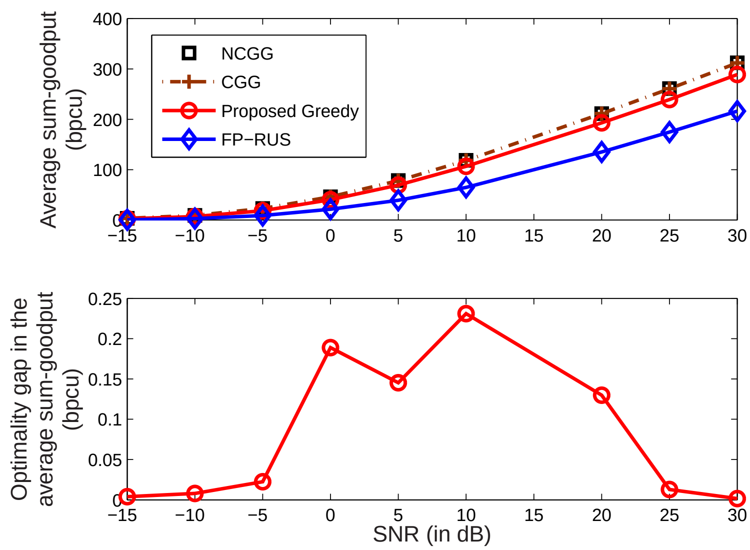

Fig. 7. The top plot shows the average sum-goodput as a function of SNR. The bottom plot shows the average bound on the optimality gap between the proposed and optimal greedy solutions (given in (30)), i.e., the average value of $(\mu^* - \mu_{\min})(X_{\mathsf{con}} - X^*_{\mathsf{tot}}(\boldsymbol{I}^{\min}, \mu^*))$. In this plot, $N = 32$, $K = 8$, $\alpha = 10^{-3}$, and $S = 30$.